PM-85

THE CRITICAL INCLINATION IN SATELLITE THEORY

ELECTRONICS RESEARCH CENTER

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

PM-85

THE CRITICAL INCLINATION IN SATELLITE THEORY

By Diarmuid O'Mathuna
NASA Electronics Research Center
Cambridge, Massachusetts

July 1969

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

INTRODUCTION

In Celestial Mechanics the occurence of small divisors has such a long tradition that their appearance, even in the absence of an underlying physical explanation, is not considered surprising (cf. e.g., Brouwer (Ref. 1)). In the case of satellite theory when certain perturbation theories in terms of instantaneous elements are applied, the small divisors arise for parameter values near the so-called critical inclination.

The perturbation of Hamiltonian systems has long been associated with the analysis of area-preserving mappings which are perturbations of twist mappings of an annulus. The celebrated result of Moser (Ref. 2) gives the conditions for the existence of invariant curves for such mappings. The conditions exclude the rational rotation numbers corresponding to the resonance cases of the system. The latter cases, with their associated questions of stability are reflected in the small divisors of the series expansions. In this respect the small divisors may reflect the physical phenomenon as well as series expansions permit.

However, since in satellite theory the small divisors arise in a manner which does not reflect the physical situation, it puts in evidence an unsatisfactory feature of perturbation theory. The present note is aimed at illustrating how this feature may arise. It appears that the problems of perturbation theory near the critical inclination stem from applying the implicit function theorem near a point where the conditions for its validity are not satisfied.

We shall consider the Vinti Model for satellite theory, which has the advantage of integrability so that an exact solution can be written. For comparison the corresponding formulae for the Kepler problem are also exhibited. In order to interpret the solution of the Vinti problem in terms of instantaneous Kepler elements it is necessary to make expansions in terms of the small parameter. It becomes evident that such expansions require particular attention near the resonant case – that is in the region of coincidence of the two basis frequencies: to a first approximation this corresponds to the so-called critical inclination. A closer examination reveals that an expansion in this region may hide an invalid application of the binomial theorem.

1. THE VINTI PROBLEM*

The Vinti problem is formulated in terms of spheroidal co-ordinates R, σ and ϕ, with spheroidal constant c. In terms of these the cartesian co-ordinates are given by the relations

$$\left.\begin{aligned} x &= (R^2 + c^2)^{\frac{1}{2}} \sin\sigma \cos\phi \\ y &= (R^2 + c^2)^{\frac{1}{2}} \sin\sigma \sin\phi \\ z &= R\cos\sigma \end{aligned}\right\} \quad (1.1)$$

so that R and σ are related to the spherical co-ordinates r and θ by the relations

$$r^2 = R^2 + c^2 \sin^2\sigma, \quad r\cos\theta = R\cos\sigma \qquad (1.2)$$

* The results stated in this section are derived in Ref. 3.

In the spheroidal system the Vinti potential has the form

$$V = -\frac{\mu R}{R^2 + c^2 \cos^2\sigma} \tag{1.3}$$

The integration of the associated dynamical problem yields the following three first integrals, namely

$$\left.\begin{aligned}
&\frac{1}{2}\,\frac{(R^2 + c^2\cos^2\sigma)^2}{R^2 + c^2}\,\dot{R}^2 + \alpha^2R^2 - \mu R = \tfrac{1}{2}\lambda_3^2\,\frac{c^2}{R^2 + c^2} - \tfrac{1}{2}\lambda^2\\
&\tfrac{1}{2}(R^2 + c^2\cos^2\sigma)\,\dot{\sigma}^2 + \alpha^2c^2\cos^2\sigma = -\tfrac{1}{2}\lambda_3^2\,\frac{1}{\sin^2\sigma} + \tfrac{1}{2}\lambda^2\\
&(R^2 + c^2)\,\sin^2\sigma\,.\,\dot{\phi} = \lambda_3
\end{aligned}\right\} \tag{1.4}$$

where the dot denotes differentiation with respect to time and the three constants λ_3, λ and α^2 are introduced by the integration. The constants λ_3 and α^2 denote the polar component of angular momentum and the negative of the energy respectively. The constant λ has the dimension of angular momentum but does not have an obvious physical interpretation except in the degenerate case when $c = 0$: in that case λ denotes the magnitude of the angular momentum vector for what is then Keplerian motion.

In terms of the above constants we can define the fundamental length scales a_o and p_o by setting

$$a_o = \frac{\mu}{2\alpha^2}\,,\quad p_o = \frac{\lambda^2}{\mu} \tag{1.5}$$

and introduce the fundamental parameters by the relations

$$\nu = \frac{\lambda_3}{\lambda}, \quad \ell_o^2 = \frac{P_o}{a_o}, \quad \eta = \frac{c}{P_o} \tag{1.6}$$

Since the constants λ_3, λ and α^2 are readily determined from initial conditions by means of the integrated relation (1.4) it follows that the fundamental length scales (1.5) and the fundamental parameters (1.6) are also immediately determined in terms of the initial data. The relevance of the problem to satellite theory is in the range

$$0 < \eta << 1 \tag{1.7}$$

The final integration of equations (1.4) is achieved by the introduction of the new independent variable f, defined by

$$\frac{df}{dt} = \frac{\Lambda}{R^2 + c^2 \cos^2 \alpha} \tag{1.8}$$

Here Λ is a constant whose ratio to λ is an algebraic form in terms of the fundamental parameters: thus Λ also has the dimension of angular momentums and and f is dimensionless. In the limit when $\eta \to o$ we would have $\Lambda \to \lambda$ and f would then become the true anomaly of the Kepler problem.

In terms of f the solution to the first pair of equations (1.4) can be represented by means of the Jacobian elliptic functions as follows

$$\left. \begin{aligned} u &= \frac{1}{R} = \frac{1}{P} \left[\frac{1 + e \,\mathrm{cn}\left[j_1 f, k_1\right]}{1 + \eta^2 d\, e\, \mathrm{cn}\left[j_1 f, k_1\right]} \right] \\ \cos \sigma &= \sqrt{1 - N^2}\ \mathrm{sn}\left[f + \omega, k_2\right] \end{aligned} \right\} \tag{1.9}$$

In the above the quantities e, N and d are algebraically related to the fundamental parameters as are also the quantities j_1, k_1, k_2 and the ratio P/P_o.

The relevant range of interest for satellite theory has been noted in (1.7); in this range we have

$$\left.\begin{aligned} j_1 &= 1 + O(\eta^2),\ k_1 = O(\eta^2),\ k_2 = O(\eta^2) \\ e^2 &= 1 - \ell_o^2 + O(\eta^2),\ 1 - N^2 = (1 - \nu^2)\left[1 + O(\eta^2)\right] \\ p &= p_o\left[1 + O(\eta^2)\right] \end{aligned}\right\} \quad (1.10)$$

The arbitrary constant ω has been introduced by the integration and we have chosen the origin for f to coincide with a minimum value for R: thus ω represents minus the value of f at the equatorial crossing prior to that minimum. In the limit when $\eta \to 0$, then ω represents the argument of perigee of the Kepler problem.

2. THE CORRESPONDING KEPLER PROBLEM

The analogous formulae for the Kepler problem can be obtained by setting c = 0 and hence also $\eta = 0$ in the foregoing. In that case $R \to r$, $\sigma \to \theta$ so that the co-ordinate system is spherical and we have

$$\left.\begin{aligned} x &= r \sin\theta \cos\phi \\ y &= r \sin\theta \sin\phi \\ z &= r \cos\theta \end{aligned}\right\} \quad (2.1)$$

and the potential is given by

$$V_K = -\frac{\mu}{r} \quad (2.2)$$

Letting $\lambda \to \lambda_K$ and $\alpha \to \alpha_K$ then corresponding to relations (1.4) we have the three well-known first integrals for the Kepler problem, namely

$$\left.\begin{aligned} &\tfrac{1}{2}\, r^2 \dot{r}^2 + \alpha^2 r^2 - \mu r = -\tfrac{1}{2}\lambda_K^2 \\ &\tfrac{1}{2}\, r^2 \dot{\theta}^2 + \tfrac{1}{2}.\lambda_3^2 \cdot \frac{1}{\sin^2\theta} = \tfrac{1}{2}\lambda_K^2 \\ &r^2 \sin^2\theta .\ \dot{\phi} = \lambda_3 \end{aligned}\right\} \qquad (2.3)$$

The semi major axis a_K and the semilatus rectum p_K are defined by

$$a_K = \frac{\mu}{a_K^2}\ ,\ P_K = \frac{\lambda_K^2}{\mu} \qquad (2.4)$$

Then the Kepler elements of inclination and eccentricity are given respectively by

$$\nu_K = \frac{\lambda_3}{\lambda_K},\ e_K^2 = 1 - \left(\frac{P_K}{a_K}\right)^2 \qquad (2.5)$$

For the solution of equation (2.3) we define the true anomaly f_k by the relation

$$\frac{df_K}{dt} = \frac{\lambda_K}{r^2} \qquad (2.6)$$

and the solutions to the first pair of equaltion (2.3) take the form

$$\left.\begin{aligned} u_K &= \frac{1}{r} = \frac{1}{P_K}\left[1 + e_K \cos f_K\right] \\ \cos\theta &= \sqrt{1 - \nu_K^2}\ \sin\left[f_K + \omega_K\right] \end{aligned}\right\} \qquad (2.7)$$

where ω_K is a constant introduced by the integration. The angle f_K is to be measured from perigee so that ω_K represents the angle of perigee. The solutions (2.7) could have been derived simply by setting $\eta = o$ in relations (1.9).

3. INSTANTANEOUS KEPLER ELEMENTS

The standard perturbation theories for satellite theory are developed for slowly varying Kepler elements. Accordingly we can get some insight into the latent difficulties if we attempt to interpret the exact solution for the Vinti problem in terms of instantaneous Kepler elements.

For this purpose we note from (1.2) that

$$\cos\theta = \frac{R}{r}\cdot\cos\sigma = \frac{R}{\sqrt{R^2 + c^2\sin^2\sigma}}\cdot\cos\sigma = \frac{1}{\sqrt{1 + \frac{c^2}{R^2}(1-\cos^2\sigma)}}\cos\sigma \tag{3.1}$$

If we now substitute for R and $\cos\sigma$ in terms of f from (1.9) and set

$$\varepsilon = \frac{c}{p} \tag{3.2}$$

we have

$$\cos\theta = \frac{\sqrt{1-N^2}\,\mathrm{sn}\left[f+\omega,k_2\right]}{\sqrt{1+\varepsilon^2\left[\frac{1+e\,\mathrm{cn}\left[j_1 f,k_1\right]}{1+\eta^2 ed\,\mathrm{cn}\left[j_1 f,k_1\right]}\right]^2\left[\mathrm{cn}^2\left[f+\omega,k_2\right]+N^2\mathrm{sn}^2\left[f+\omega,k_2\right]\right]}} \tag{3.3}$$

In term of instantaneous Kepler elements we would have

$$\cos\theta = \sqrt{1-\nu_K^{\,2}}\ \sin\ (f_K + \omega_K) \tag{3.4}$$

The determination of the slowly varying Kepler elements follows from identification of the representations (3.3) and (3.4). Both represent the same oscillatory motion; by identifying the amplitudes we obtain the formula for the element ν_K; while the identification of the residual normalized oscillation gives the formula for the element ω_K.

Proceeding in this manner we obtain for ν_K the relation

$$\nu_K = \sqrt{1 - \frac{1 - N^2}{1 + \varepsilon^2 \left[\frac{1+e\ \mathrm{cn}\left[j_1 f, k_1\right]}{1+\eta^2 \mathrm{de}\ \mathrm{cn}\left[j_1 f, k_1\right]}\right]^2 \left[\mathrm{cn}^2\left[f+\omega, k_2\right] + N^2 \mathrm{sn}^2\left[f+\omega, k_2\right]\right]}} \tag{3.5}$$

while from the normalized oscillation there follows

$$\sin\,(f_K + \omega_K) = \mathrm{sn}\left[f + \omega, k_2\right] = \sin\left(\mathrm{am}\left[f + \omega, k_2\right]\right) \tag{3.6}$$

from which we immediately get for ω_K

$$\omega_K = \mathrm{am}\left[f + \omega,\ k_2\right] - f_K \tag{3.7}$$

The next step is to determine f in term of f_K which when introduced into relations (3.5) and (3.7) would yield the appropriate representations for the instantaneous elements ν_K and ω_K.

For the remainder we shall confine our attention to the formula for ω_K; in this we shall show that if we seek an approximate representation by making a series expansion in terms of the small parameter, particular attention is necessary in the range near the critical inclination: in fact the general validity of the series expansion appears questionable in that range.

4. PERTURBATION EXPANSION FOR THE ARGUMENT OF PERIGEE

In obtaining the perturbation series we first note that the frequency associated with the Jacobian elliptic functions is an analytic function of the modulus. In fact for the modulus k_2 we have the series expansion valid for k_2 near zero,

$$\frac{\pi}{2K_2} = 1 - \frac{1}{4}k_2^2 - \frac{5}{64}k_2^4 + \ldots \tag{4.1}$$

where $4K_2$ is associated period. Also for the modulus q_2 of the associated Theta functions we have

$$q_2 = \frac{1}{16}k_2^2\left[1 + \frac{1}{2}k_2^2 + \frac{21}{64}k_2^4 + \ldots\right] \tag{4.2}$$

Then with G defined by

$$G = \frac{\pi}{2K_2}(f + \omega) \tag{4.3}$$

We have the following trigonometric series representation for the am function appearing in relation (3.7), namely

$$\mathrm{am}\left[f+\omega,\ k_2\right] = G + \sum_{n=1}^{\infty}\frac{2q_2^n}{n(1+q_2^{2n})}\cos 2nG \tag{4.4}$$

so that from (3.7) we have

$$\omega_\kappa = \frac{\pi}{2K_2}\omega + \left(f-f_\kappa\right) + \left(\frac{\pi}{2K_2} - 1\right)f + \sum_{n=1}^{\infty}\frac{2q_2^\eta}{n(1+q_2^{2n})}\cos 2nG \tag{4.5}$$

It remains to determine the perturbation series for f in terms of f_κ: this is done by first obtaining the representation for f_κ as a perturbation series in terms of f and thence obtaining the inverse relationship. To obtain the expansion for f_κ in terms of f we combine equation (1.8) with equation (3.6) to give

$$\frac{df_k}{df} = \frac{\lambda_\kappa}{\Lambda}\frac{R^2 + c^2\cos^2\sigma}{r^2} = \frac{\lambda_\kappa}{\Lambda}\frac{R^2 + c^2\cos^2\sigma}{R^2 + c^2\sin^2\sigma}$$

$$= \frac{\lambda_\kappa}{\Lambda}\frac{1 + \left(c^2/R^2\right)\cos^2\sigma}{1 + \left(c^2/R^2\right)\sin^2\sigma} = \frac{\lambda_\kappa}{\Lambda}\frac{1 + \varepsilon^2(pu)^2\cos^2\sigma}{1 + \varepsilon^2(pu)^2\sin^2\sigma}$$

$$= \frac{\lambda_\kappa}{\Lambda}\left[1 + \varepsilon^2(pu)^2\cos^2\sigma\right]\left\{1 - \varepsilon^2(pu)^2\sin^2\sigma + \varepsilon^4(pu)^4\sin^4\sigma + \ldots\right\} \tag{4.6}$$

where the terms not explicitly written are of order η^6. If we multiply out the product on the right and re-arrange we obtain

$$\frac{df_\kappa}{df} = \frac{\lambda_\kappa}{\Lambda} \Big\{ 1 + \varepsilon^2 (pu)^2 (2\cos^2\sigma - 1) + \varepsilon^4 (pu)^4 (1 - \cos^2\sigma)(1 - 2\cos^2\sigma) + \ldots \Big\} \tag{4.7}$$

The next step is to substitute for u and $\cos\sigma$ in terms of f from (1.9) and then perform the integration to the desired degree of accuracy. The integration leads to both secular and periodic terms: the character of the terms is unambiguously determined except for those terms resulting from products of the form

$$\mathrm{cn}\left[j_1 f, k_1\right] \cdot \mathrm{cn}\left[f + \omega, k_2\right] \tag{4.8}$$

that is, for terms reflecting the direct interaction of the basic oscillations of the problem.

The character of the terms resulting from the integration of such combinations depends on whether or not we are dealing with case of resonance, i.e. whether*

$$\left.\begin{array}{lll} \text{(i)} & \dfrac{K_1}{j_1} = K_2 & \text{(R)} \\ \text{or} & & \\ \text{(ii)} & \dfrac{K_1}{j_1} \neq K_2 & \text{(NR)} \end{array}\right\} \tag{4.9}$$

In case (1) there arises a secular term, while in case (ii) there arise a low frequency term with the small divisor

$$\frac{K_1}{j_1} - K_2 \equiv \frac{K_2}{j_1}\left(\frac{K_1}{K_2} - j_1\right) \tag{4.10}$$

introduced by the integration. This indicates that in and near the Resonance case (R) one must proceed cautiously in interpreting the

*$4K_1$ is the period of the elliptic function with modulus k_1.

solution in terms of perturbed instantaneous Kepler elements.

If one calculates the factors in the divisor (4.10) as a perturbation series in η^2 we find that

$$\frac{K_2}{j_1} = 2\pi \left[1 + O(\eta^2)\right] , \quad \frac{K_1}{K_2} - j_i = O(\eta^2) \tag{4.11}$$

and in fact we have

$$\frac{K_1}{j_1} - K_2 = -\frac{3\pi}{2}\eta^2 \left[\alpha_2 + \eta^2 \alpha_4 + \eta^4\alpha_6 + \ldots\right] \tag{4.12}$$

where in terms of the fundamental parameters

$$\alpha_2 = 5\nu^2 - 1 \tag{4.13}$$

Now the procedure of perturbation theory requires a series expansion for the reciprocal of the divisor (4.10); this is obtained by making a binomial expansion for the reciprocal of the series (4.12) in the form

$$\begin{aligned}\left[\frac{K_1}{j_1} - K_2\right]^{-1} &= -\frac{2}{3\pi}\frac{1}{\eta^2\alpha_2}\left[1 + \eta^2\frac{\alpha_4}{\alpha_2} + \eta^4\frac{\alpha_6}{\alpha_2} + ..\right]^{-1} \\ &= -\frac{2}{3\pi}\cdot\frac{1}{\eta^2\alpha_2}\left[1 + \eta^2\frac{\alpha_4}{\alpha_2} + \eta^4\left(\frac{\alpha_4^2}{\alpha_2^2} - \frac{\alpha_6}{\alpha_2}\right) + \ldots\right]\end{aligned} \tag{4.14}$$

Unless we are dealing with the particular case where α_2 divides all α_{2n} for $n > 2$, such an expansion can be valid at best for

$$\alpha_2 > \eta^2 \tag{4.15}$$

More specifically we must in general exclude the range near $\alpha_2 = 0$, that is the range near

$$5\nu^2 - 1 = 0 \tag{4.16}$$

To a first approximation this corresponds to the well-known critical inclination.

CONCLUSION

The separability of the Vinti dynamical problem allows the solution to be written in the exact form (1.9), giving the two coordinates R and cos σ in terms of the Vinti anomaly f. In this form each of the coordinates has associated with it a specific frequency namely the frequency of the respective elliptic function appearing on the right hand side of (1.9). In general these frequencies are distinct. In the degenerate case ($\eta = 0$) the solution becomes the solution for the Kepler problem and the frequencies coincide. In the non-degenerate case ($\eta \neq 0$) there is coincidence only for a preferred set of initial parameters: to first order in η^2 this preferred set corresponds to what is called the critical inclination in the perturbation methods of satellite theory.

The representation (1.9) is valid for all parameter values: in particular, no difficulties arise in the region of the preferred set mentioned above. However, if one attempts to convert this form of solution into a representation for the instantaneous Kepler elements in terms of the instantaneous true anomaly the resulting form may not be uniformly valid for all parameter values. Particular attention must then be given to the preferred set where difficulties may arise in the determination of the series expansions.

The difficulties exhibit themselves in the relation between the Vinti anomaly f and the instantaneous true anomaly f_K: near the critical inclination the validity of a series expansion is doubtful. The terms giving rise to the difficulty stem from the direct interaction of terms involving the two (in general) distinct frequencies mentioned above. The interaction produces "resonant" denominators near the coincidence of the two frequencies: the series expansion requires the expansion of the reciprocal of such denominators; near the critical inclination the validity of such an expansion is at least questionable.

It should also be noted that associated with the preferred set is the occurrence of periodic solutions: thus for this set we must expect that $d\omega_K/df_K = 0$. The vanishing of this derivative makes questionable the application of the implicit function theorem to the relation (3.7) in a range that includes the critical inclination: the problem arising in the series expansion is but a reflection of this difficulty.

This difficulty arising in the series representation for Kepler elements is in clear contrast to the representation (1.9). In the latter the forms in terms of f are separated so that no interaction occurs. This emphasizes an often ignored feature of separability: apart from allowing the integrability of the problem in terms of the appropriate independent variable, it also permits the solution to the written in such a form that the basic frequencies do not interact. If one retains the form suggested by the separation, one need not anticipate any difficulty in the region of coincidence of the frequencies.

REFERENCES

1. Brouwer, D., and Clemence, G. M.: Methods of Celestial Mechanics, Academic Press, New York, 1962.

2. Moser, J.: On Invariant Curves of Area Preserving Mappings of an Annulus. Nachr. Akad. Wiss. Göttingen, Math. Phys. Kl, IIa, no. 1, 1962, pp. 1-20.

3. O'Mathuna, D.: The Vinti Dynamical Problem and the Geopotential. NASA TR R-307, ERC, Cambridge, Mass., 1969.